# Health-Conscious Integrated Thermal Management Strategy Using Hybrid Attention Deep Reinforcement Learning for Battery Electric Vehicles

Changcheng Wu, Jiankun Peng, Dawei Pi, Xin Guo, Hailong Zhang, Zexing Wang, and Yunpeng Li

***Abstract*—Effective thermal management strategies (TMSs) can extend the driving range of battery electric vehicles (BEVs) and enhance cabin thermal comfort in high-temperature environments. Considering the development trend toward integrated thermal management systems (ITMSs) and the critical role of the power battery in BEVs, this article establishes an ITMS model embedded with battery health-conscious. To further explore the temperature control and energy-saving potential of the proposed ITMS, a learning TMS is designed using twin delayed deep deterministic policy gradient (TD3). Given the intricate state information within the ITMS, a hybrid attention mechanism is introduced to optimize the original TD3, enabling the TD3 agent to discern the relative importance of various state information and thereby improving its decision-making quality. Simulation results indicate that, compared to rule-based TMS, the proposed improved TD3-based TMS decrease battery health degradation and energy loss by 22.50% and 35.33%, respectively. Additionally, compared to other compared TMSs, the proposed TMS effectively reduces temperature fluctuations in the battery, motor, and cabin, with root mean square errors of 0.13, 0.12, and 0.08, respectively. Moreover, hardware in the loop testing further demonstrates that the proposed TMS can reduce driving costs by 9.13%–12.70% compared to the traditional TMS.**

***Index Terms*—Battery health-conscious, hybrid attention mechanism, integrated thermal management systems (ITMSs), thermal management strategies (TMSs), twin delayed deep deterministic policy gradient.**

Received 24 March 2025; revised 22 May 2025; accepted 18 June 2025. Date of publication 20 June 2025; date of current version 5 August 2025. This work was supported in part by the China's National Key R&D Program under Grant 2022YFB4300300, in part by the National Natural Science Foundation of China under Grant 52372380, in part by the National Natural Science Foundation of China under Grant 52402522, and in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant SJCX25_0124. Recommended for publication by Associate Editor H. Chaoui. *(Corresponding author: Jiankun Peng.)*

Changcheng Wu, Jiankun Peng, Xin Guo, and Yunpeng Li are with the School of Transportation, Southeast University, Nanjing 211189, China (e-mail: 230238983@seu.edu.cn; jkpeng@seu.edu.cn; guoxin023@seu.edu.cn; 220243656@seu.edu.cn).

Dawei Pi is with the School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: pidawei@njust.edu.cn).

Hailong Zhang is with the College of Mechanics Engineering, North University of China, Taiyuan 030051, China (e-mail: 20220143@nuc.edu.cn).

Zexing Wang is with National New Energy Vehicle Technology Innovation Center, Beijing 100048, China (e-mail: wangzexing@nevc.com.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2025.3581978.

Digital Object Identifier 10.1109/TPEL.2025.3581978

## Nomenclature

| | |
|---|---|
| BEVs | Battery electric vehicles. |
| AC | Air condition. |
| BTMSs | Battery thermal management systems. |
| ITMSs | Integrated thermal management systems. |
| TMSs | Thermal management strategies. |
| MPC | Model predictive control. |
| DP | Dynamic programming. |
| PTC | Positive temperature coefficient. |
| DRL | Deep reinforcement learning. |
| DQN | Deep Q-network. |
| DDPG | Deep deterministic policy gradient. |
| CTMC | Cabin thermal management circuit. |
| BTMC | Battery thermal management circuit. |
| ACTMC | Air conditioning thermal management circuit. |
| MTMC | Motor thermal management circuit. |
| TD3 | Twin delayed deep deterministic policy gradient. |
| SOH | State of health. |
| EOL | End of life. |
| WLTC | World light vehicle test cycle. |
| HWFET | Highway fuel economy test. |
| UDDS | Urban dynamometer driving schedule. |
| NEDC | New European driving cycle. |
| RSME | Root mean square error. |
| HIL | Hardware in the loop. |
| VCU | Vehicle control unit. |

## I. Introduction

BATTERY electric vehicles (BEVs) are highly acclaimed by governments worldwide due to their environmental friendliness and nonpolluting characteristics [1]. However, there are numerous factors constraining their rapid development. In high-temperature conditions, the air condition (AC) systems of BEVs can reduce the driving range by approximately 18%–37% [2], hindering the process of replacing traditional fuel vehicles.

Thermal management systems are a key research focus in academia and industry, addressing the essential safety, economy, comfort, and durability needs of BEVs. In the early stages of development for BEVs, the thermal management function of the battery, motor, and cabin systems operated independently [3].

Lithium-ion batteries are widely favored in energy storage applications due to their superior energy density and extended cycle life [4]. However, their operational stability under high-current discharge conditions is constrained by three critical challenges: thermal accumulation; accelerated capacity fade; and potential thermal runaway risks [5]. To address these multiphysics challenges, recent advancements in battery thermal management systems have emerged through dual optimization approaches: novel pack configurations incorporating phase-change materials and microchannel cooling [6], [7], and development of hybrid composite materials with anisotropic thermal conductivity and electrochemical stability [8], [9]. Liu et al. [10] compared the impact of different condenser pipe layouts on the cooling performance of the battery system. Yang et al. [11] improved the cooling performance of the motor system by reducing the height and width of the liquid-cooled pipes. Khayyam et al. [12] dynamically adjusted the blower speed and recirculation-air gate opening in the cabin, achieving a 14% reduction in energy consumption of air conditioning while ensuring thermal comfort. Although distributed thermal management systems are structurally simple and easy to control, they suffer from drawbacks such as a large number of components, significant volume and mass, and high energy consumption [13].

To address these issues, leading international vehicle technology research and development institutions are focusing their efforts on the development of more compact, energy-efficient, and wider temperature range integrated thermal management systems (ITMSs) [14], [15]. With the increased integration of the ITMSs, the complexity and difficulty of its control strategies have also increased correspondingly [16], which intensifies the pursuit of more intelligent, efficient, and integrated thermal management strategies (TMSs). Researchers have been tirelessly working towards this goal. Hong et al. [17] proposed a rule-based integrated TMS that extended the driving range of BEV by 24.2% in summer and 18.6% in winter. However, rule-based TMSs are often constrained by engineers' experience, which can prevent them from achieving optimal thermal management performance. Guo et al. [18] found that compared to the rule-based integrated TMS, a model predictive control (MPC)-based integrated TMS reduced energy consumption by 2.84%. Hajidavalloo et al. [19] introduced a battery-cabin integrated TMS using nonlinear MPC, which successfully improved the vehicle's driving range while enhancing cabin comfort. Lian et al. [20] proposed an integrated TMS based on dynamic programming, which improved the vehicle's economy and comfort by optimizing the control trajectories of the compressor speed, power of positive temperature coefficient, and waste heat mode. However, the TMSs based on these optimization theories face significant computational burdens, leading to poor real-time performance and hindering their large-scale real-world application.

As artificial intelligence technology advances rapidly, deep reinforcement learning (DRL)-based TMSs have developed into an effective solution for optimally balancing the two types of TMSs. Choi et al. [21] indicated that compared to rule-based integrated TMS, the total energy consumption of integrated TMS based on deep Q-network (DQN) is reduced by approximately 15.6%. Zhang and Huang [22] designed a cooling control strategy based on deep deterministic policy gradient (DDPG) for battery and cabin systems, demonstrating a 7.3% reduction in compressor energy consumption compared to PID-based control strategy. Wei et al. [23] compared TMSs based on optimization theory and DDPG, revealing that the latter outperforms the former in terms of power battery efficiency and overall vehicle economy. However, the literature survey reveals several research gaps: First, while there has been significant progress in DRL-based TMSs, their application is often limited to one or two thermal circuits [24], [25]. Their modeling paradigms employed struggle to represent the electro-mechanical-thermal dynamic coupling mechanisms within multiloop ITMSs (involving battery-motor-cabin-AC), which constrains the capability for cross-domain energy coordination optimization of ITMSs. Additionally, current research primarily focuses on maintaining optimal temperature ranges for internal components and reducing energy consumption. However, in BEVs, few studies consider battery health status in the optimization metrics for TMSs, despite the high cost of battery replacement [26]. Finally, previous studies based on DRL [21], [22], employing a homogeneous state-action weight allocation strategy, exhibit a deficiency in dynamically identifying critical information (such as overheating and aging risks) when handling multiloop thermal management collaborative control state inputs. This limitation results in the control strategy's inability to adaptively suppress the interference of secondary objectives (e.g., cabin temperature fluctuations) on core safety indicators under long-term driving conditions.

To bridge the identified research gaps, the three contributions of the current study are follows:

1) Based on the battery-motor-cabin coupling model and TD3 algorithm, a cross-domain thermal management cooperative control system for BEVs is proposed to achieve dynamic optimal energy allocation of multiple heat sources.
2) This article presents a pioneering approach to incorporating battery health-conscious into an integrated TMS, which mitigates the costly degradation of power battery and enhances the overall performance and reliability of BEVs.
3) Giving the intricate thermodynamic characteristics within the ITMS, a hybrid attention mechanism is incorporated to enhance the DRL agent's ability to focus on critical information, thereby elevating the decision-making quality of the TMS.

The rest of this article is as follows. Section II delineates the development of the ITMS model. Section III presents an improved TD3 based TMS. Section IV is dedicated to the discussion and analysis of the experimental results. Finally, Section V concludes this article.

## II. Mathematical Modeling of ITMS

### A. Architecture of ITMS

The ITMS consists of a cabin thermal management circuit (CTMC), a battery thermal management circuit (BTMC), a

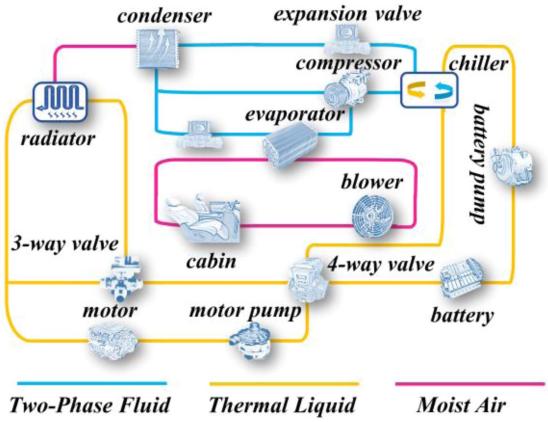


Fig. 1. Architecture of ITMS.

motor thermal management circuit (MTMC) and an air conditioning thermal management circuit (ACTMC), as shown in Fig. 1. The BTMC contains the following components: battery; battery pump; and chiller. The MTMC contains the following components: motor; motor pump; and radiator. The BTMC and MTMC are connected via a four-way valve. This article only investigates the cooling performance of the ITMS in hot weather. The four-way valve connects BTMC and MTMC in parallel under this assumption. The three-way valve determines whether the radiator is involved in the motor's cooling process. The CTMC consists of the cabin, blower and evaporator. The ACTMC consists of the compressor, expansion valve, condenser, evaporator and chiller. The ACTMC cools the battery and cabin by transferring heat through the phase change of refrigerant in the refrigeration cycle. CTMC and ACTMC are connected through the evaporator, which is used to provide air conditioning in hot weather. ATCMC is linked to MTMC and BTMC via radiator and chiller, respectively. When the radiator alone is inadequate, the chiller is employed to cool the coolant during hot weather. This model is developed based on the Simscape library in MATLAB/Simulink [27].

It is noteworthy that the ITMS model is built under the following assumptions:

1) Pump
   a) Pump is considered as a quasi-steady component.
   b) The pump wall is rigid.
   c) Ignoring external leaks from the pump.
2) Compressor
   a) The compressor does not accumulate mass.
   b) Flow is quasi-steady within compressor.
3) Blower
   a) There are no irreversible losses.
   b) There is no heat exchange with the environment.
4) Fan
   a) Fan is considered as a quasi-steady component.
   b) Use static pressure rise rather than total fan pressure to simulate fan performance.
5) Pipe
   a) The pipe material is isotropic.
   b) Ignore temperature changes and deformations in pipe.

### B. Motor Thermal Model

Both motors are synchronous motors with permanent magnets. This article simplifies thermal model of the motor and its heat generation $Q_{\text{mot}g}$ is directly calculated based on the working power $P_{\text{mot}}$ and working efficiency $\eta_{\text{mot}}$ [28]

$$\begin{cases} Q_{\text{mot}g} = P_{\text{mot}}\left(1-\eta_{\text{mot}}\right) \\ P_{\text{mot}} = v\left(m_v g f + \frac{C_D A}{21.15} v^2 + \delta m_v \dot{v}\right) \end{cases} \tag{1}$$

where $v$ and $m_v$ are speed and mass of selected vehicle, respectively; $f$ and $C_D$ are rolling resistance and aerodynamic drag coefficients, respectively; $g$ is the gravitational acceleration; $A$ is the frontal area; $\delta$ is the rotational mass conversion coefficient.

According to the first law of thermodynamics, the temperature rises $\Delta T_{\text{mot}}$ of the motor can be calculated by the following equation [29]:

$$\begin{cases} c_{\text{mot}} m_{\text{mot}} \Delta T_{\text{mot}} = Q_{\text{mot}g} - Q_{\text{mot}d} \\ Q_{\text{mot}d} = h_m A_m \left(T_{\text{mot}} - T_{m\_c}\right) \end{cases} \tag{2}$$

where $c_{\text{mot}}$ denotes the specific heat capacity of the motor, $m_{\text{mot}}$ denotes the mass of the motor; $Q_{\text{mot}d}$ is the heat dissipation of motor; $h_m$ and $A_m$ denote the heat transfer coefficient and heat exchange area between the motor and the coolant, respectively; and $T_{m\_c}$ is the temperature of motor coolant.

### C. Battery Thermal Model With Health Conscious

The investigated battery cell operates at a rated voltage of 3.3 V with an energy storage capacity of 7.6 Wh. Its electrode configuration consists of a graphite anode paired with a lithium iron phosphate cathode, while the electrolyte material employs lithium hexafluorophosphate. The power demand $P_{\text{bat}}$ of the power battery while driving BEV can be expressed as (3) [30]

$$\begin{cases} P_{\text{bat}} = P_{dr} + P_{\text{ITMS}} \\ P_{dr} = \begin{cases} \frac{P_{\text{mot}}}{\eta_{\text{mot}}}, \dot{v} \geq 0 \\ P_{\text{mot}} \eta_{\text{mot}}, \dot{v} < 0 \end{cases} \end{cases} \tag{3}$$

where $P_{\text{ITMS}}$ is the total energy consumption of the ITMS, which is the sum of the energy consumption of the compressor, fan, blower, battery pump and motor pump.

According to reference [31], heat generation $Q_{\text{bat}g}$ of the power battery can be expressed as

$$\begin{cases} Q_{\text{bat}g} = I_{\text{bat}}^2 R_{\text{bat}} + I_{\text{bat}} T_{\text{bat}} \frac{dU_{\text{bat}}}{dT_{\text{bat}}} \\ I_{\text{bat}} = \frac{U_{\text{oc}} - \sqrt{U_{\text{oc}}^2 - 4 R_{\text{bat}} P_{\text{bat}}}}{2 R_{\text{bat}}} \\ U_{\text{bat}} = U_{oc} - I_{\text{bat}} R_{\text{bat}} \end{cases} \tag{4}$$

where $I_{\text{bat}}$ is battery current; $U_{\text{oc}}$ is battery open-circuit voltage; $R_{\text{bat}}$ is battery internal resistance; $T_{\text{bat}}$ is battery temperature; and $U_{\text{bat}}$ is battery terminal voltage.

According to [30], the battery temperature variation $\Delta T_{\text{mot}}$ can be modelled as the following equation:

$$\begin{cases} c_{\text{bat}} m_{\text{bat}} \Delta T_{\text{bat}} = Q_{\text{bat}g} - Q_{\text{bat}d} \\ Q_{\text{bat}d} = h_b A_b \left(T_{\text{bat}} - T_{b\_c}\right) \end{cases} \tag{5}$$

where $c_{\text{bat}}$ is specific heat capacity of battery; $m_{\text{bat}}$ is the mass of battery; $h_b$ is convective heat transfer coefficient; $A_b$ is the surface area of battery; $Q_{\text{bat}d}$ is the heat dissipation of battery; and $T_{b\_c}$ is the temperature of battery coolant.

Given the high cost of battery, its aging is taken into account. The capacity loss $\Delta Q_{\text{bat}}$ of the battery can be defined as follows [32]:

$$\begin{cases} \Delta Q_{\text{bat}} = B(c) \cdot e^{\frac{-E_a(c)}{8.31 T_{\text{bat}}}} \cdot Ah(c)^{0.55} \\ E_a(c) = 31700 - 370.3 \cdot c \end{cases} \tag{6}$$

where $B$ is the pre-exponential factor; $c$ is the C-rate, which is calculated based on reference [33]; Ah is the total ampere-hour throughput; and $E_a$ denotes activation energy.

The capacity loss of the battery at its end of life (EOL) is usually considered to be 20%, so the $\text{Ah}(c,\ T_{\text{bat}})$ and the total number of cycles $N(c,\ T_{\text{bat}})$ before reaching its EOL can be calculated as follows [34]:

$$\begin{cases} \text{Ah}(c, T_{\text{bat}}) = \left( \frac{20}{B(c) \cdot \exp\left( \frac{-E_a(c)}{8.31 T_{\text{bat}}} \right)} \right)^{\frac{1}{0.55}} \\ N(c, T_{\text{bat}}) = \frac{3600 Ah(c, T_{\text{bat}})}{Q_{\text{bat}}} \end{cases}. \tag{7}$$

The decline in the state of health (SOH) of battery under multiple factors is given by the following equation [35]:

$$\frac{d\text{SOH}}{dt} = -\frac{1}{2N(c, T_{\text{bat}}) Q_{\text{bat}}} \int_0^t |I_{\text{bat}}(\tau)|\, d\tau. \tag{8}$$

### D. Cabin Model

The cabin thermal load originates from both outer and inner sources. Externally, load is transferred into the cabin via thermal conduction and solar radiation. Internally, the load includes heat from occupants, the ventilation system, and the AC system [36]. The vehicle body material's heat transfer properties cause heat transfer between the cabin interior and the external environment, which can be expressed as follows [31]:

$$Q_{ie} = \sum_{i=1}^{4} h_{\text{sc}} A_i (T_{\text{env}} - T_{\text{cab}}) \tag{9}$$

where $Q_{ie}$ is conduction thermal load; $h_{\text{sc}}$ is the coefficient of convective heat transfer; $A_i$ are heat transfer area of the roof panel, windshield, rear window and side window, respectively; and $T_{\text{env}}$ and $T_{\text{cab}}$ are temperature of environment and cabin, respectively.

Solar radiation increases the heat load, which can be expressed as follows [31]:

$$Q_{sr} = \sum_{i=1}^{4} \chi S_r G_i f_s sin\sigma_i \tag{10}$$

where $Q_{sr}$ is radiation thermal load; $\chi$ is the transmissivity of solar radiation by glasses; $S_r$ is the solar radiation intensity; $f_s$ is shading factor; $G_i$ are the area of the windshield, rear windows, side windows, roof panel, respectively; $\sigma_i$ are angles between windshield, rear windows, side windows, roof panel and the vertical direction.

This article assumes that there is only one driver in the vehicle, whose heat generation is empirically about 145 W [37]. When the AC system is turned on, the ventilation system circulates air inside and outside the cabin, potentially causing internal heat generation or loss. This thermal load of ventilation system $Q_{vs}$ is modelled as follows [37]:

$$Q_{vs} = m_{vs} \xi c_{ca} (T_{\text{env}} - T_{\text{cab}}) \tag{11}$$

where $m_{vs}$ is the mass of airflow; $\xi$ is air recirculation coefficient; and $c_{ca}$ is specific heat capacity of the cabin air.

Based on the above equations, the temperature change of cabin can be expressed as follows [31]:

$$\Delta T_{\text{cab}} = \frac{(Q_{ie} + Q_{sr} + Q_{vs} + 145\ W \cdot s) - Q_{\text{com}}}{\rho_{ca} V_{ca} c_{ca}} \tag{12}$$

where $\rho_{ca}$ is the density of cabin air; $V_{ca}$ is the volume of cabin air; and $Q_{\text{com}}$ is the cooling capacity of AC.

### E. Models of Compressor, Fan, Pump and Blower

According to reference [38], the mechanical efficiency $\eta_{\text{mec}}$ and motor efficiency $\eta_{c\_m}$ of the compressor can be calculated by (13). Then, according to these efficiencies, the compressor power $P_{\text{com}}$ can be expressed by (14)

$$\begin{cases} \eta_{\text{mec}} = 0.8680 + 0.0048 N_{\text{com}} - 4.4444 * 10^{-5} N_{\text{com}}^2 \\ \eta_{c_m} = 0.6980 + 0.0013 N_{\text{com}} + 4.1235 * 10^{-5} N_{\text{com}}^2 \\ -4.8781 * 10^{-7} N_{\text{com}}^3 + 1.4206 * 10^{-9} N_{\text{com}}^4 \end{cases} \tag{13}$$

$$\begin{cases} h_{\text{dis}} = h_{\text{suc}} + \frac{h_{\text{dis}}^* - h_{\text{suc}}}{\eta_{\text{isen}}} \\ m_{\text{com}} = N_{\text{com}} \rho_{\text{suc}} \left( V * 10^{-6} \right) \\ P_{\text{com}} = \frac{W_{\text{com}} (h_{\text{dis}} - h_{\text{suc}})}{\eta_{\text{mec}} \eta_{c\_m}} \end{cases} \tag{14}$$

where $N_{\text{com}}$ is compressor speed; $h_{\text{dis}}$, $h_{\text{suc}}$, and $h_{\text{dis}}^*$ are discharge specific enthalpy, suction specific enthalpy, and isentropic discharge specific enthalpy, respectively; $\eta_{\text{isen}}$ denotes isentropic efficiency; $W_{\text{com}}$ is compressor flow rate; $\rho_{\text{suc}}$ is suction density; and $V$ is compressor displacement.

The fan operation process is modelled as follows [28]:

$$\begin{cases} P_{\text{fan}} = T_{\text{fan}} w_{\text{fan}} \\ T_{\text{fan}} = \Phi_{\text{ref}}(q_{\text{ref}}) \frac{w_{\text{fan}}^2}{w_{\text{fanRef}}^3} \left( \frac{\rho}{\rho_{\text{ref}}} \right) \left( \frac{D_{\text{fan}}}{D_{\text{fanRef}}} \right)^5 \\ q_{\text{ref}} = q \frac{w_{\text{fanRef}}}{w_{\text{fan}}} \left( \frac{D_{\text{ref}}}{D} \right)^3 \end{cases} \tag{15}$$

where $P_{\text{fan}}$ is the power of fan; $T_{\text{fan}}$ and $w_{\text{fan}}$ are torque and angular velocity of the fan shaft, respectively; $w_{\text{fanRef}}$ is reference shaft speed; $\rho$ and $\rho_{\text{ref}}$ are the moist air density and reference density, respectively; $D_{\text{fan}}$ and $D_{\text{fanRef}}$ are fan diameter and reference diameter, respectively; $q$ and $q_{\text{ref}}$ are volumetric flow rate and reference volumetric flow rate, respectively; $\Phi_{\text{ref}}$ is the proportion of the pressure gain to the fan efficiency, which is modeled as the function of $q_{\text{ref}}$ in this article.

According to [39], the power consumption $P_{\text{pump}}$ of pump is shown as follows:

$$\begin{cases} P_{\text{pump}} = \frac{W_{\text{pump}} \Delta p_{\text{pump}}}{\eta_{\text{pump}} \rho_c} \\ \Delta p_{\text{pump}} = 0.927 W_{\text{pump}}^2 + 0.586 W_{\text{pump}} - 0.143 \end{cases} \tag{16}$$

where $W_{\text{pump}}$ is pump flow rate; $\Delta p_{\text{pump}p}$ is the press drop of the pump; $\eta_{\text{pump}}$ is pump efficiency; and $\rho_c$ is the coolant density.

According to reference [40], the blower is modelled as follows:

$$P_{bl} = \beta_1 W_{bl}^2 + \beta_2 W_{bl} + \beta_3 \tag{17}$$

where $P_{bl}$ is the power of blower; $W_{bl}$ is blower flow rate; $\beta_1$, $\beta_2$ and $\beta_3$ are the fitted parameters, their values are [21456 −1974.2 49.318].

## III. DRL-Based TMS

### A. Theory of TD3

The TD3 is an improved version of the DDPG created to boost its performance and stability. It employs three techniques to enhance DDPG [32], [41].

1) TD3 employs two critic-networks to evaluate the policy and substitutes the smaller of their outputs into the bellman equation to calculate the loss function:

$$\begin{cases} y_t = r_t + \gamma \min\limits_{i=1,2} \left(Q'_{\theta_i}\left(s'_{t+1}, a'_{t+1}\right)\right) \\ J\left(\theta_i\right) = \min\limits_{\theta_i} E_\pi \left[\left(y_t - Q_{\theta_i}\left(s_t, a_t\right)\right)^2\right], i = 1, 2 \end{cases} \tag{18}$$

where $y_t$ is the target Q-value in a certain state $s_t$ and action $a_t$; $r_t$ is the instantaneous reward; $Q_\theta$ and $Q'_\theta$ are critic-network and target critic-network, respectively; $\gamma$ is discount factor; $J$ is the loss function; $E$ is mathematical expectation.

2) Adding clipped normal distributed noise to the output actions of the target actor-network to eliminate overfitting

$$a' = \pi'_\Phi\left(s'\right) + \varepsilon, \varepsilon \sim c_{\text{lip}}\left(N\left(0, \sigma\right), -c, c\right), c > 0 \tag{19}$$

where $\pi_\Phi{}'$ is target actor-network and $\varepsilon$ is the clipped noise.

3) To enhance TD3 stability, the parameters' update of actor and target networks are less frequent. Updates happen only after the critic-networks have been updated for a set number of steps. The updating process of the actor-critic networks are determined by the following equation [41]:

$$\begin{cases} \nabla_{\theta_i} J\left(\theta_i\right) = E_{\pi_\Phi} \\ \left[\left(y_t - Q_{\theta_i}\left(s_t, a_t\right)\right)^2 \nabla_{\theta_i} Q_{\theta_i}\left(s_t, a_t\right)\right], i = 1, 2 \\ \nabla_\Phi J\left(\Phi\right) = E_{\pi_\Phi} \\ \left[\nabla_a Q_{\theta_i}\left(s_t, a_t\right)|_{a=\pi_\Phi(s)} \nabla_\Phi \pi_\Phi\left(s_t\right)\right] \end{cases} \tag{20}$$

where, $\nabla_a Q_{\theta_i}(s_t, a_t)$ is gradient of Q-value at $a_t$; $\nabla_\Phi \pi_\Phi(s_t)$ is the gradient of the policy; $\nabla_{\theta_i} J(\theta_i)$ is the gradient of loss function; $\nabla_\Phi J(\Phi)$ is the gradient of target function; $\nabla_{\theta_i} Q_{\theta_i}(s_t, a_t)$ is the gradient of Q-value at $\theta_i$.

The updating process of the target actor-critic networks are determined by the following equation [41]:

$$\begin{cases} \theta'_i \leftarrow \tau\theta_i + (1-\tau)\,\theta'_i, i = 1, 2 \\ \Phi' \leftarrow \tau\Phi + (1-\tau)\,\Phi' \end{cases} \tag{21}$$

where $\theta'$ and $\Phi'$ are the parameters of target critic and target actor networks, respectively; $\theta$ and $\Phi$ are the parameters of critic and actor networks, respectively; $\tau$ is the soft updating factor.

### B. Specific Implements

The decision-making process of DRL involves three key elements, i.e., the state $s$, the action $a$ and the reward function $R$

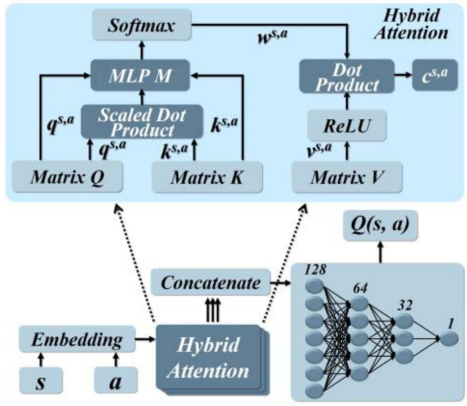


Fig. 2. Structure of improved critic network.

[42]. The state $s$ must comprehensively represent crucial environmental features, allowing agents to make informed decisions. The state $s$ is defined as

$$s = [T_{\text{cabin}}, T_{\text{mot}}, T_{\text{bat}}, P_{\text{ITMS}}, O_{\text{val3}}, O_{\text{val4}}] \tag{22}$$

where $T_{\text{cabin}}$ is cabin temperature; $T_{\text{mot}}$ is motor temperature; $T_{\text{bat}}$ is battery temperature; $P_{\text{ITMS}}$ is the total energy consumption of the ITMS; $O_{\text{val3}}$ and $O_{\text{val4}}$ are the operating states of the three-way and four-way valves, respectively.

The operating states of the 3-way and 4-way valves, compressor speed $N_{\text{com}}$, fan speed $N_{\text{fan}}$, blower speed $N_{\text{blo}}$, motor pump speed $N_{\text{mot\_pump}}$ and battery pump speed $N_{\text{bat\_pump}}$ are selected as the action. Details on operation of all valves are available in our previous study [28]

$$a = [O_{\text{val3}}, O_{\text{val4}}, N_{\text{com}}, N_{\text{fan}}, N_{\text{blo}}, N_{\text{mot\_pump}}, N_{\text{bat\_pump}}]. \tag{23}$$

This article aims to stabilize the internal temperature of the ITMS while minimizing battery health decline and its energy consumption, thus the reward function is designed as

$$\begin{aligned} R = & -k_1\left(T_{\text{cabin}} - T_{cf}\right) - k_2\left(T_{\text{mot}} - T_{mf}\right) \\ & - k_3\left(T_{\text{bat}} - T_{bf}\right) - k_4 P_{\text{ITMS}} - k_5 \Delta \text{SOH}_b \end{aligned} \tag{24}$$

where $T_{cf}$ is the reference temperature of cabin (25 °C); $T_{mf}$ is the reference temperature of motor (60 °C); $T_{bf}$ is the reference temperature of battery (30 °C); $\Delta\text{SOH}_b$ is the battery health decline; $k_i (i = 1, 2, \ldots, 5)$ are user-defined weights for balancing various optimization objectives.

### C. TD3 in Combination With Hybrid Attention

State reflects current information and action predicts future information. To learn the optimal decision-making strategy, state and action need to be given different attention levels so as to extract the most critical current and future information [43].

For the above purpose, this section embeds a hybrid attention mechanism [44] based on key-value pairs into TD3

$$q^{s,a} \rightarrow \begin{cases} q^s = Qe_i\left(s\right) \\ q^a = Qn_i\left(a\right) \end{cases} \tag{25}$$

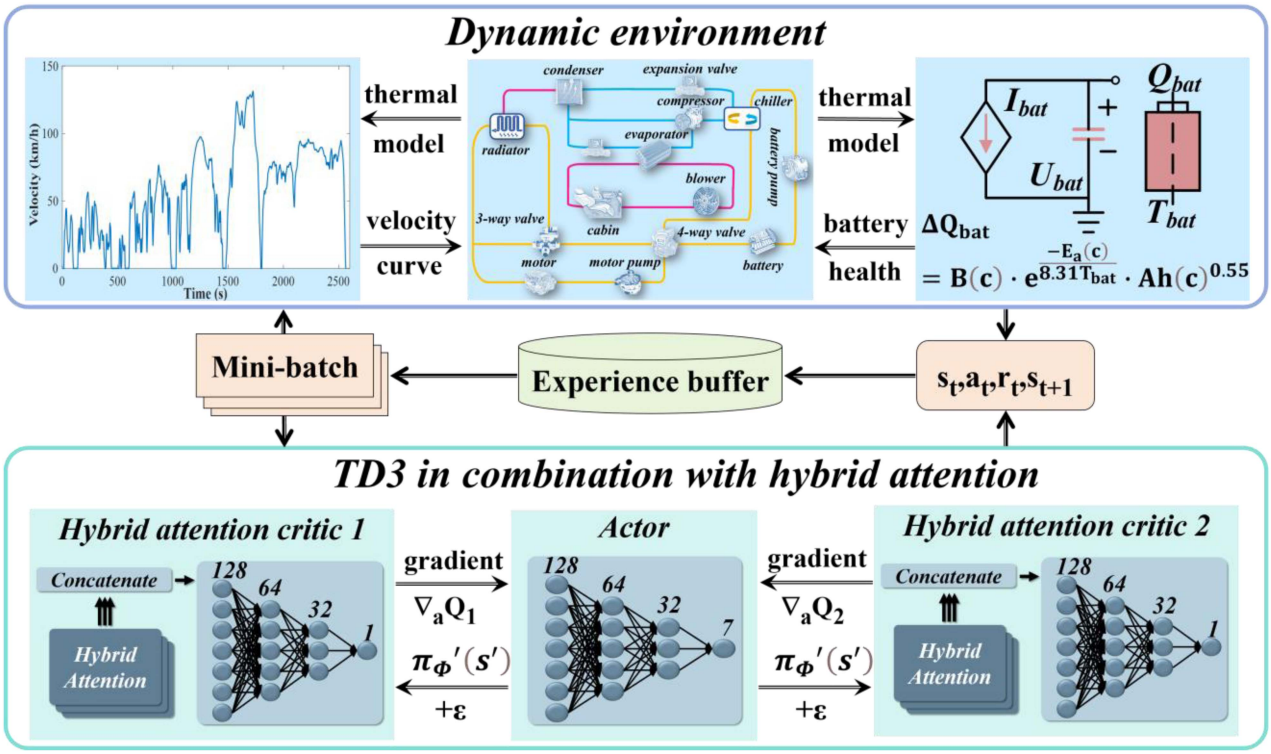


Fig. 3. Schematic diagram of an enhanced TD3-based TMS.

$$k^{s,a} \rightarrow \begin{cases} k^s = Ke_i(s) \\ k^a = Kn_i(a) \end{cases}. \tag{26}$$

In the proposed mechanism, the state and action of the TMS are encoded as dense representations. The state-action value function is represented by $Q_i(s,a) = f_i(e_i(s), n_i(a))$. The $f_i$ is a two-layer multilayer perceptron, consisting of a single-layer perceptron $e_i$ for the state and another single-layer perceptron $n_i$ for the action.

The hybrid attention mechanism is depicted in topmost submodule of Fig. 2. The shared matrix $V$ linearly transforms the embedded code to produce the value function $v^{s,a}$ based on the state and action, respectively; $w^{s,a}$ reflects the attention to the state and action; The shared matrices $Q$ and $K$ linearly transforming the embedded code into their respective query $q^{s,a}$ and key $k^{s,a}$, as shown in (25) and (26); Attention is determined by the similarity between the query and the key. The similarity is calculated by an additional multilayer perceptron $M$, and the calculation result is passed to SoftMax

$$w^{s,a} = \text{SoftMax}\left(M\left(\frac{q^{s,a}k^{s,a}}{\sqrt{d_k}}\right), q^{s,a}, k^{s,a}\right) \tag{27}$$

where $\sqrt{d_k}$ denotes scaling factor.

The critic network embedded with the hybrid attention mechanism is shown in Fig. 2. The improved critic network's information processing capacity is significantly boosted by the differential attention allocation weight for the state and action. Additionally, the multi-head hybrid attention enhances the model's scalability. The contributions of each head can be concatenated as follows:

$$c_i^x = \text{Concatenate}\left(c_{i1}^x, c_{i1}^x, \ldots, c_{ih}^x\right) \tag{28}$$

where $x$ refers to the state $s$ or the action $a$, and $h$ is the number of heads.

Finally, the design of the proposed TD3-based TMS is depicted in Fig. 3. Additionally, the hyperparameters for TD3, as established through multiple simulations, are given in Table I. The value limitations of action are given in Table II.

TABLE I
HYPERPARAMETER SETTINGS

| Parameters | Value |
|---|---|
| Hidden layers (Actor and Critic networks) | 3,3 |
| Number of neurons per hidden layer | 128, 64, 32 |
| Number of neurons per input layer | 6,13 |
| Number of neurons per output layer | 7,1 |
| Optimizer | Adam |
| Discount factor | 0.99 |
| Learning rate for Critic network | 5e-5 |
| Learning rate for Actor network | 5e-4 |
| Exploration noise | 0.15 |
| Experience replay buffer size | 1e6 |
| Minibatch size | 64 |
| Training episodes | 800 |

## IV. TRAINING, COMPARISONS, AND DISCUSSIONS

In the automotive sector, the world light vehicle test cycle and the highway fuel economy test serve as standardized test cycles,

TABLE II
VALUE LIMITATIONS OF THE ACTION

| Notations | Description | Value |
|---|---|---|
| $N_{\text{fan}}$ | Fan command (r/min) | [0 3000] |
| $N_{\text{comp}}$ | Compressor command (r/min) | [0 3500] |
| $N_{\text{blo}}$ | Blower command ($m^3/s$) | [0 0.15] |
| $N_{\text{bat\_pump}}$ | Battery pump command (r/min) | [0 1000] |
| $N_{\text{mot\_pump}}$ | Battery pump command (r/min) | [0 1000] |
| $O_{val4}$ | Four-way valve command | 0, 1 |
| $O_{val3}$ | Three-way valve command | 0, 1 |

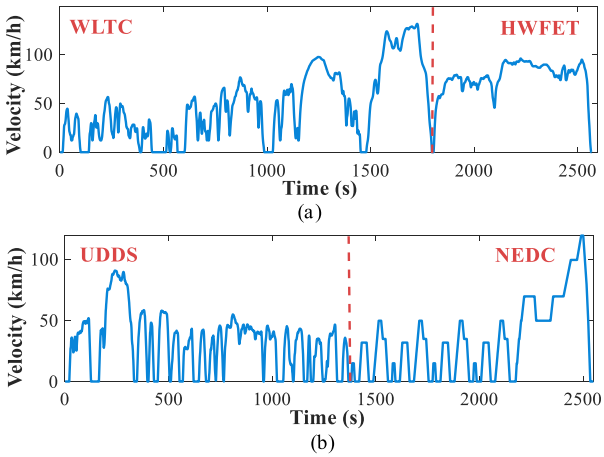


Fig. 4. Velocity profiles for training and test cycles. (a) Velocity profile for training cycle. (b) Velocity profile for test cycle.

each originating from diverse traffic and driving scenarios. This article synthesizes these cycles to create the training cycle that encompasses a broad spectrum of speeds and encompasses both urban and suburban environment. To evaluate the adaptability of the proposed TMS under new driving conditions, this article integrates the urban dynamometer driving schedule with the new European driving cycle to form a composite test cycle. Training and test cycles are shown in Fig. 4.

### A. Analysis of Optimal Performance

This section evaluates the training performance, a key measure of the proposed DRL-based TMS, i.e., TMS@A. To validate the effectiveness of TD3 in ITMS, research methods from reference [21] (i.e., DQN-based TMS, TMS@N) and [22] (i.e., DDPG-based TMS, TMS@G) are used for comparison. Fig. 5(a) illustrates the mean reward of episodes, where the proposed TMS@A outperforms TMS@G and TMS@N in terms of both convergence speed and convergence reward. The underlying reason for this phenomenon may be attributed to the fact that, compared to DDPG and DQN, TD3 effectively mitigates the risk of converging to local optima through its dual critic network architecture.

The temperature variations of the battery, motor, and cabin are plotted in Fig. 5(b), (c), and (d). It shows that proposed TMS@A can be effectively kept at their target temperature

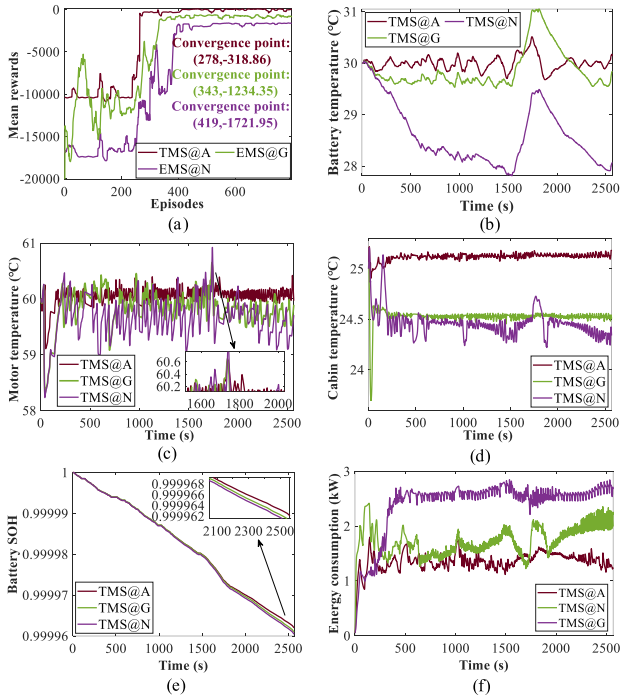


Fig. 5. Performance analysis under the training cycle. (a) Learning curves. (b) Battery temperature curves. (c) Motor temperature curves. (d) Cabin temperature curves. (e) Battery SOH trajectorie. (f) Power variation curves.

during operation. Fig. 5(e) shows that battery health declines over time, where battery health degradation of the proposed TMS@A is consistently lower than the other two compared TMSs, and its final value is about 0.999963. Fig. 5(f) shows that the proposed TMS can maintain the power in the range of 1.5 kW. However, the TMS@G has a power range of up to 2.5 kW, which shows the significant energy saving potential of the proposed TMS@A.

### B. Analysis of Adaptability Performance

Taking into account that the suggested TMS is to be integrated into the vehicle end, and that the vehicle is subject to diverse driving conditions, it is crucial to assess the adaptability of the proposed TMS across these varying conditions. Fig. 6 illustrates the temperature fluctuations of the battery, motor, and cabin under various TMSs.

Rule-based TMS, i.e., TMS@R, results in wider temperature fluctuation ranges for controlled components due to his inability to dynamically adapt to the complex and variable thermal dynamics of the ITMS. This is compounded by the inherent limitations of predefined rules in capturing the nuanced thermal behavior under a diverse driving condition. In contrast, the proposed TMS@A significantly mitigates the temperature fluctuations across the battery, motor, and cabin. This is attributed to TMS@A's ability to learn generalized features of the environment, enabling adaptive refinement of the control strategy to ensure robust adaptability. The TMS based on MPC i.e., TMS@M, exhibits inferior stability in temperature regulation compared to the proposed TMS@A. This discrepancy

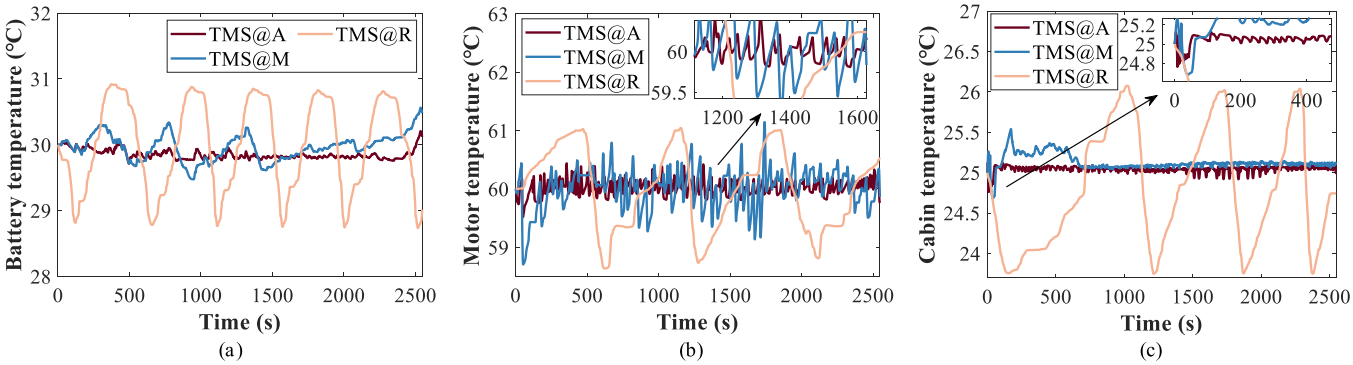


Fig. 6. Verification of temperature curves under the test cycle. (a) Battery temperature curves. (b) Motor temperature curves. (c) Cabin temperature curves.

TABLE III
HEALTH DECLINE AND ENERGY CONSUMPTION OF BATTERY FOR VARIOUS TMSS UNDER THE TEST CYCLE

| Categories | Battery health decline | Improvement | Total energy consumption (kWh) | Improvement |
|---|---|---|---|---|
| TMS@R | 4.0e-5 | - | 1.33 | - (baseline) |
| TMS@M | 3.6e-5 | 10.00% | 1.15 | 13.53% |
| TMS@A | 3.1e-5 | 22.50% | 0.86 | 35.33% |

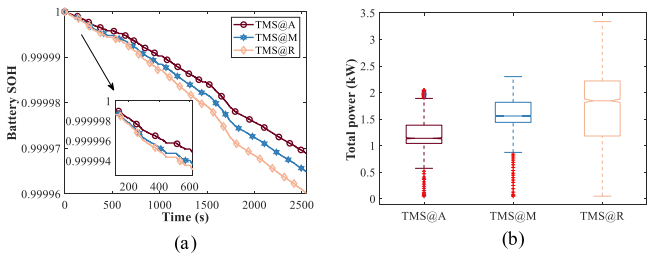


Fig. 7. Performance analysis under the training cycle. (a) Battery SOH trajectories. (b) Variation of total power.

is primarily attributed to the dependency of MPC on the accuracy of the predefined model. In contrast, TD3 algorithms, as utilized in TMS@A, continuously learn and refine control strategies through interaction with the environment, thereby demonstrating enhanced adaptability to model uncertainties and environmental fluctuations.

The decline trajectories of battery health for various TMSs are shown in Fig. 7(a). Battery SOH value of TMS@R decline from 1 to about 0.9980 due to its inability to include battery health awareness in its optimization metrics. TMS@A and TMS@M embedded with battery health awareness significantly slows down battery health degradation compared to TMS@R. Moreover, TMS@A outperforms TMS@M with a final battery SOH value of approximately 0.9991. The total power of the ITMS for three different TMSs is drawn in Fig. 7(b). The TMS@A presents the lowest median, which suggests that the proposed method requires lower operating power in most cases. In addition, the wide ranges of the upper and lower quartiles in the TMS@R indicate substantial power fluctuations. Table III gives a quantitative comparison of the phenomena in Fig. 7(a) under the test cycle. TMS@A and TMS@M scale reduce battery health degradation by 10.00% and 22.50%, respectively, compared to TMS@R. Moreover, TMS@A and TMS@M decrease total energy consumption by 13.53% and 35.33%, respectively, compared to TMS@R. The above results show that the proposed TMS@A possesses excellent adaptability under new driving conditions.

The temperature control capability and overall energy consumption of the ITMS are significantly influenced by auxiliary components (i.e., compressor, fan, and blower). Therefore, Fig. 8 analyzes the density distribution of these components during operation. Fig. 8(a) illustrates that the compressor predominantly operates at approximately 400 r/min under TMS@A and TMS@M, whereas TMS@R primarily operates around 600 r/min. Obviously, the compressors of TMS@A and TMS@M consume less energy. Fig. 8(b) indicates that the fan speed of TMS@A is primarily in the low-speed range. TMS@M exhibits a broader distribution of fan speed, and TMS@R operates at higher speeds, reflecting a greater operational load. Fig. 8(c) illustrates that the blower flow rate of TMS@A is predominantly concentrated in the low-flow region. In contrast, the blower flow rate of TMS@M exhibits a relatively broader distribution with an increase in flow rates. The TMS@R towards the high-flow area, indicating a higher load on the blower under this strategy.

### C. Ablation Experiments

This section sets comparative TMSs for ablation experiment of the health awareness model and hybrid attention mechanism proposed in this article, with specific settings given in Table IV. Performance analysis of comparative TMSs is presented in Table V. The TMS@A that considers battery health and hybrid attention converges more slowly, taking 176.5 min, as the agent must balance multiple optimization objectives. However, this complexity enables the proposed TMS@A to achieve better control. All TMSs outperform TMS@R in mitigating battery health decline. Among the all compared TMSs, TMS@A and TMS@B show the healthiest battery states, improving by 22.50% and 20.00%, respectively, based on TMS@R, due to

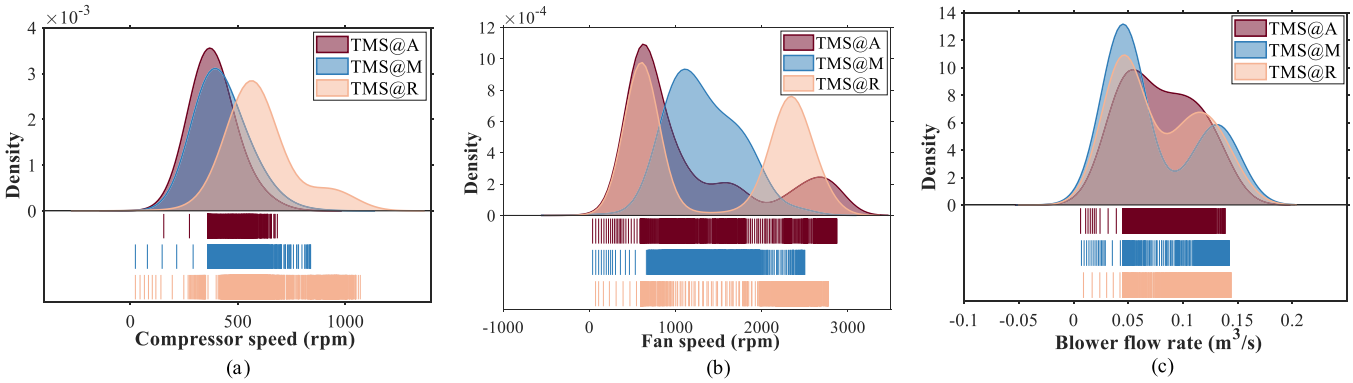


Fig. 8. Distribution maps of the components' work. (a) Distribution of compressor work. (b) Distribution of fan work. (c) Distribution of blower work.

TABLE IV
SETTINGS OF VARIOUS TMSS FOR ABLATION EXPERIMENT

| Categories | Algorithms | Battery health | Hybrid attention |
|---|---|---|---|
| TMS@A | TD3 | ● | ● |
| TMS@B | TD3 | ● | ○ |
| TMS@C | TD3 | ○ | ● |
| TMS@D | TD3 | ○ | ○ |

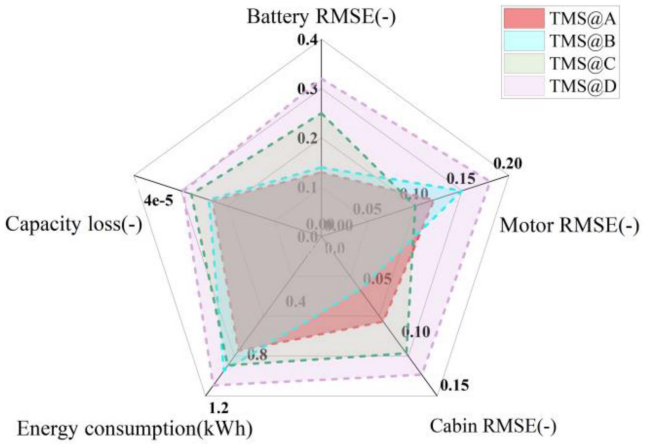


Fig. 9. Distribution maps.

the integration of health awareness in their learning processes. TMS@D without embedding battery health awareness achieves a 1.67% health improvement over TMS@R, likely due to TD3's trial-and-error learning ability. All TMSs show improvements in total energy consumption, with TMS@A leading at a reduction of 35.33%, followed by TMS@C at 27.04%. TMS@D presents a smaller improvement of 15.79%. The possible reason for this phenomenon is that the hybrid attention mechanism improves the quality of decision-making by giving agents the ability to distinguish the importance of state information, thereby promoting the emergence of optimal solutions.

Fig. 9 shows the performance of four TMSs across five key metrics. Overall, TMS@A is the best performer in all metrics except for motor and cabin temperature root mean square error (RSME) among all TMSs. Its battery temperature RSME is about 0.12, battery capacity loss is about 0.0009 and energy

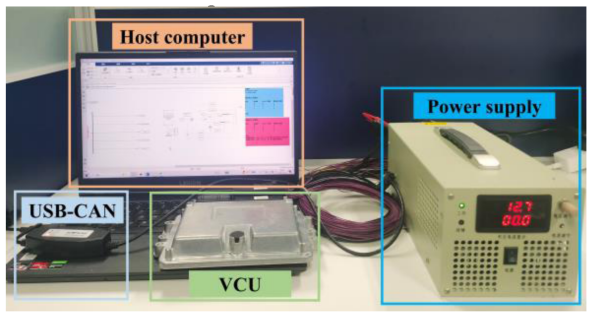


Fig. 10. Configuration of HIL system.

consumption is 0.86 kWh, indicating high temperature control accuracy, battery SOH and energy efficiency. TMS@C performs the best on motor temperature RSME because it ignores battery health state and overly focuses on motor temperature stability with the help of hybrid attention mechanism. TMS@B achieves the highest performance in cabin temperature RMSE. This likely results from unprioritized state information, causing excessive battery health focus and compressor overuse that over-restrict cabin heating.

### D. Analysis of HIL Test Results

Hardware in the loop (HIL) testing is essential for validating the performance and reliability of the proposed TMS on actual hardware, ensuring accurate implementation and optimal performance in real vehicle. The HIL system is illustrated in Fig. 10, which includes the host computer (Windows 11, Intel i7-13700 CPU, and RTX3060 GPU), the USB-CAN (Ecotrons CAN), the vehicle control unit (VCU) [VCU22297A03, Infineon TC297], and power supply (12 V). Specifically, the ITMS model is constructed by Matlab/Simulink (Version 2023b) platform. Then well-trained TMS@A is flashed into the VCU. Finally, the VCU communicates with the ITMS model through the USB-CAN to realize the expected action control of components. The control command cycle of the VCU with baud rate 250 k is 10 ms, and the running time of the proposed TMS@A is 7.8 ms.

Moreover, to analyze the sensitivity of the proposed TMS@A to initial temperatures, this section defines two distinct cases: case A, where the initial temperatures of both the battery and

TABLE V
PERFORMANCE ANALYSIS OF COMPARATIVE TMSS

| Categories | Convergence time (min) | Battery health decline | Improvement | Energy consumption (kWh) | Improvement |
|---|---|---|---|---|---|
| TMS@R | - | 4.0e−5 | - (baseline) | 1.33 | - (baseline) |
| TMS@A | 176.5 | 3.1e−5 | 22.50% | 0.86 | 35.33% |
| TMS@B | 125.6 | 3.2e−5 | 20.00% | 1.01 | 24.06% |
| TMS@C | 152.3 | 3.7e−5 | 7.50% | 0.97 | 27.04% |
| TMS@D | 115.7 | 3.9e−5 | 1.67% | 1.12 | 15.79% |

TABLE VI
TOTAL DRIVING COSTS FOR HIL TEST

| Cases | Categories | Battery aging cost (CNY) | Battery charging cost (CNY) | Total Cost (CNY) | Improvement |
|---|---|---|---|---|---|
| Case A | TMS@R | 12.98 | 3.77 | 16.75 | - (baseline) |
| | TMS@M | 12.48 | 3.56 | 16.04 | 4.23% |
| | TMS@A (S) | 11.75 | 3.34 | 15.09 | 9.91% |
| | TMS@A (H) | 11.84 | 3.38 | 15.22 | 9.13% |
| Case B | TMS@R | 6.13 | 1.43 | 7.56 | - (baseline) |
| | TMS@M | 5.79 | 1.31 | 7.10 | 6.08% |
| | TMS@A (S) | 5.43 | 1.07 | 6.50 | 14.02% |
| | TMS@A (H) | 5.49 | 1.11 | 6.60 | 12.70% |

Note: S denotes simulation results; H denotes HIL results.

motor are 45 °C, the initial cabin temperature is 30 °C, and the ambient temperature is 32 °C; and case B, where the initial battery temperature is 25 °C, the initial motor temperature is 75 °C, the initial cabin temperature is 20 °C, and the ambient temperature is 35 °C. A traffic cycle based on real data [45], illustrated in Fig. 11(a), is used to better align the HIL test with actual driving conditions. The monitoring results depicted in Fig. 11(b)–(g) indicate that the proposed TMS@A within the actual vehicle control unit operates as intended, effectively maintaining the temperatures of the battery, motor, and cabin within the desired ranges. This phenomenon demonstrates that the proposed TMS@A possesses the ability to operate stably and improve the performance of ITMS in real vehicles.

The data in Table VI reveal that proposed TMS@A significantly reduce total driving costs compared to typical TMS@R and TMS@M that are widely used. Compared to TMS@R, in case A, TMS@A achieves a 9.13% cost reduction, while in case B, the improvement increases to 12.70%. The primary driver of cost savings is the reduction in battery aging costs, highlighting the effectiveness of optimized TMS@A in enhancing battery lifespan and economic efficiency under varying operating conditions. Moreover, the HIL test performance of the proposed TMS@A differs from the simulation test performance by 0.78% and 1.32% in both cases, demonstrating good robustness.

To further validate the dominance of the proposed TMS under different driving cycles, Fig. 12 statistics the total driving costs of different TMSs under federal test procedure 75, China light-duty vehicle test cycle for passenger cars, Japan cycle 08 and world harmonized vehicle cycle. These comparisons are built on a fair environment where they have the same ambient and component initial temperatures (i.e., 45 °C for motor and battery, 30 °C for cabin, 32 °C for ambient). TMS@A exhibits the least cost-effectiveness under the JC08 cycle, with a cost of 5.51 CNY. Moreover, compared to other TMSs, the proposed

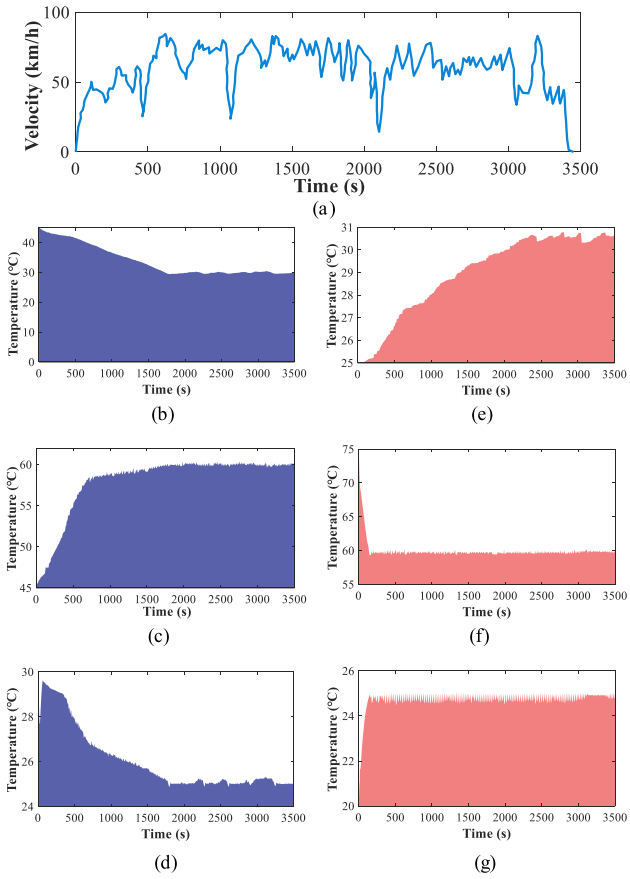


Fig. 11. Results for HIL test. (a) Cycle for HIL test. (b) Battery temperature curve for case A. (c) Motor temperature curve for case A. (d) Cabin temperature curve for case A. (e) Battery temperature curve for case B. (f) Motor temperature curve for case B. (g) Cabin temperature curve for case B.

TMS@A has the lowest driving cost under all driving cycles. This phenomenon proves that the proposed TMS@A maintains the established advantages under different driving cycles.


Fig. 12. Total driving costs under different drive cycles.

## V. Conclusion

This article develops a TD3-based battery health-conscious TMS for the ITMS in hot weather, enhancing TD3 with hybrid attention mechanism for high-quality decision-making. Based on the simulation results, the following conclusions can be further derived:

1) The convergence, the effectiveness of temperature control, energy consumption performance, and battery SOH are all indicators that demonstrate the practicality and effectiveness of the trained TMS@A under the training cycle.
2) The proposed TMS@A improves battery health degradation and energy consumption by 22.50% and 35.33%, respectively, compared to the rule-based TMS under the test cycle.
3) Ablation experiments comparing the three baselines demonstrates the effectiveness of the proposed method for battery health-conscious optimization and hybrid attention mechanism.
4) Further HIL experiments show that the proposed TMS@A saves 9.13% and 12.70% of driving costs in both cases, compared to conventional TMS@R.

While this article pioneers the integration of battery health-conscious optimization into TMS control, it currently excludes health status for other critical powertrain components, notably motor and its controller. Future work will establish a multi-objective co-optimization framework encompassing health management of battery, motor and its controller through hierarchical multi-agent RL architectures, where each agent governs specific subsystems while coordinating via attention-based meta-policies.

## References

[1] X. Zhao, X. Li, D. Jiao, Y. Mao, J. Sun, and G. Liu, "Policy incentives and electric vehicle adoption in China: From a perspective of policy mixes," *Transp. Res. Part A Policy Pract.*, vol. 190, 2024, Art. no. 104235, doi: 10.1016/j.tra.2024.104235.

[2] L. Pan, C. Liu, Z. Zhang, T. Wang, J. Shi, and J. Chen, "Energy-saving effect of utilizing recirculated air in electric vehicle air conditioning system," *Int. J. Refrigeration*, vol. 102, pp. 122–129, 2019, doi: 10.1016/j.ijrefrig.2019.03.018.

[3] J. Ma, Y. Sun, S. Zhang, J. Li, and S. Li, "Experimental study on the performance of vehicle integrated thermal management system for pure electric vehicles," *Energy Convers. Manage.*, vol. 253, Dec. 2021, Art. no. 115183, doi: 10.1016/j.enconman.2021.115183.

[4] J. H. Puet al., "Design and performance of a compact lightweight hybrid thermal management system using phase change material and liquid cooling with a honeycomb-like structure for prismatic lithium-ion batteries," *J. Power Sources*, vol. 624, 2024, Art. no. 235632, doi: 10.1016/j.jpowsour.2024.235632.

[5] W. Ma et al., "A mechanism-data driven resistance transfer algorithm for lithium-ion batteries and its application to thermal modeling," *J. Energy Storage*, vol. 102, 2024, Art. no. 114066, doi: 10.1016/j.est.2024.114066.

[6] W. Li et al., "An efficient two-stage heating strategy for embedded heat pipe system considering power and energy requirements from battery," *Appl. Therm. Eng.*, vol. 257, 2024, Art. no. 124499, doi: 10.1016/j.applthermaleng.2024.124499.

[7] S. Panchal, "Enhancing efficiency in air-cooled cylindrical battery temperature management systems for electric vehicles: A CFD analysis of a novel uniform flow distribution plate," 2024.

[8] A. Bais, D. Subhedar, and S. Panchal, "Experimental investigations of a novel phase change material and nano enhanced phase change material based passive battery thermal management system for Li-ion battery discharged at a high C rate," *J. Energy Storage*, vol. 103, 2024, Art. no. 114395, doi: 10.2139/ssrn.4743339.

[9] S. Nagaraja et al., "Influence of the fly ash material inoculants on the tensile and impact characteristics of the aluminum AA 5083/7.5 SiC composites," *Materials*, vol. 14, no. 9, 2021, Art. no. 2452.

[10] Z. Liu, X. Liu, H. Meng, L. Guo, and Z. Zhang, "Numerical analysis of the thermal performance of a liquid cooling battery module based on the gradient ratio flow velocity and gradient increment tube diameter," *Int. J. Heat Mass Transfer*, vol. 175, 2021, Art. no. 121338, doi: 10.1016/j.ijheatmasstransfer.2021.121338.

[11] C. Yang, H. Wang, X. Niu, J. Zhang, and Y. Yan, "Design and analysis of cycling oil cooling in driving motors for electric vehicle application," in *Proc. IEEE Veh. Power Propulsion Conf.*, 2016, pp. 1–6, doi: 10.1109/VPPC.2016.7791651.

[12] H. Khayyam, A. Z. Kouzani, E. J. Hu, and S. Nahavandi, "Coordinated energy management of vehicle air conditioning system," *Appl. Thermal Eng.*, vol. 31, no. 5, pp. 750–764, 2011, doi: 10.1016/j.applthermaleng.2010.10.022.

[13] L. He, H. Jing, Y. Zhang, P. Li, and Z. Gu, "Review of thermal management system for battery electric vehicle," *J. Energy Storage*, vol. 59, Nov. 2022, Art. no. 106443, doi: 10.1016/j.est.2022.106443.

[14] K. Li et al., "Assessment method of the integrated thermal management system for electric vehicles with related experimental validation," *Energy Convers. Manage.*, vol. 276, Sep. 2022, Art. no. 116571, doi: 10.1016/j.enconman.2022.116571.

[15] R. Gulfam et al., "Design, fabrication and numerical analysis of compact thermal management system integrated with composite phase change material and thermal bridge," *Energy Convers. Manage.*, vol. 156, pp. 25–33, May 2017, doi: 10.1016/j.enconman.2017.10.098.

[16] S. Jones-Jackson, R. Rodriguez, Y. Yang, L. Lopera, and A. Emadi, "Overview of current thermal management of automotive power electronics for traction purposes and future directions," *IEEE Trans. Transp. Electrific.*, vol. 8, no. 2, pp. 2412–2428, Jun. 2022, doi: 10.1109/TTE.2022.3147976.

[17] J. Hong, J. Song, U. Han, H. Kim, H. Choi, and H. Lee, "Performance investigation of electric vehicle thermal management system with thermal energy storage and waste heat recovery systems," *eTransportation*, vol. 20, Nov. 2023, Art. no. 100317, doi: 10.1016/j.etran.2024.100317.

[18] R. Guo, L. Li, Z. Sun, and X. Xue, "An integrated thermal management strategy for cabin and battery heating in range-extended electric vehicles under low-temperature conditions," *Appl. Thermal Eng.*, vol. 228, Apr. 2023, Art. no. 120502, doi: 10.1016/j.applthermaleng.2023.120502.

[19] M. R. Hajidavalloo, J. Chen, Q. Hu, Z. Song, X. Yin, and Z. Li, "NMPC-based integrated thermal management of battery and cabin for electric vehicles in cold weather conditions," *IEEE Trans. Intell. Veh.*, vol. 8, no. 9, pp. 4208–4222, Sep. 2023, doi: 10.23919/ACC55779.2023.10156539.

[20] Y. B. Lian, H. P. Ling, J. P. Zhu, J. L. Lv, and Z. W. Xie, "Thermal management optimization strategy of electric vehicle based on dynamic programming," *Control Eng. Pract.*, vol. 137, Mar. 2023, Art. no. 105562, doi: 10.1016/j.conengprac.2023.105562.

[21] W. Choi, J. W. Kim, C. Ahn, and J. Gim, "Reinforcement learning-based controller for thermal management system of electric vehicles," in *Proc. IEEE Veh. Power Propulsion Conf.*, 2022, pp. 1–5, doi: 10.1109/VPPC55846.2022.10003470.

[22] Y. Zhang and J. Huang, "Reinforcement learning-based control for the thermal management of the battery and occupant compartments of electric vehicles," *Sustain. Energy Fuels*, vol. 8, no. 3, pp. 588–603, 2024, doi: 10.1039/d3se01403g.

[23] Z. Wei, R. Song, D. Ji, Y. Wang, and F. Pan, "Hierarchical thermal management for PEM fuel cell with machine learning approach," *Appl. Thermal Eng.*, vol. 236, Feb. 2023, Art. no. 2024, doi: 10.1016/j.applthermaleng.2023.121544.
[24] G. Huang, P. Zhao, and G. Zhang, "Real-time battery thermal management for electric vehicles based on deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 14060–14072, Aug. 2022, doi: 10.1109/JIOT.2022.3145849.
[25] H. Cheng, S. Jung, and Y. B. Kim, "Battery thermal management system optimization using Deep reinforced learning algorithm," *Appl. Thermal Eng.*, vol. 236, 2024, Art. no. 121759, doi: 10.1016/j.applthermaleng.2023.121759.
[26] W. Chen, J. Peng, J. Chen, J. Zhou, Z. Wei, and C. Ma, "Health-considered energy management strategy for fuel cell hybrid electric vehicle based on improved soft actor critic algorithm adopted with Beta policy," *Energy Convers. Manage.*, vol. 292, Jul. 2023, Art. no. 117362.
[27] MathWorks, "Simscape Fluids User's Guide," 2024, [Online]. Available: https://ww2.mathworks.cn/help/pdf_doc/hydro/index.html
[28] C. Wu, J. Peng, J. Zhou, X. Guo, H. Guo, and C. Ma, "Thermal management methodology based on a hybrid deep deterministic policy gradient with memory function for battery electric vehicles in hot weather conditions," *IEEE Trans. Transp. Electrific.*, vol. 11, no. 3, pp. 7232–7242, Jun. 2025, doi: 10.1109/TTE.2024.3525014.
[29] X. Guo, J. Peng, H. He, C. Wu, H. Zhang, and C. Ma, "Integrated thermal-energy management for electric vehicles in high-temperature conditions using hierarchical reinforcement learning," *Expert Syst. Appl.*, vol. 276, Mar. 2025, Art. no. 127221, doi: 10.1016/j.eswa.2025.127221.
[30] C. Pan, Z. Jia, J. Wang, L. Wang, and J. Wu, "Optimization of liquid cooling heat dissipation control strategy for electric vehicle power batteries based on linear time-varying model predictive control," *Energy*, vol. 283, Sep. 2023, Art. no. 129099, doi: 10.1016/j.energy.2023.129099.
[31] L. Deng, S. Li, X. Tang, K. Yang, and X. Lin, "Battery thermal- and cabin comfort-aware collaborative energy management for plug-in fuel cell electric vehicles based on the soft actor-critic algorithm," *Energy Convers. Manage.*, vol. 283, Feb. 2023, Art. no. 116889, doi: 10.1016/j.enconman.2023.116889.
[32] R. Huang, H. He, X. Zhao, Y. Wang, and M. Li, "Battery health-aware and naturalistic data-driven energy management for hybrid electric bus based on TD3 deep reinforcement learning algorithm," *Appl. Energy*, vol. 321, Jan. 2022, Art. no. 119353, doi: 10.1016/j.apenergy.2022.119353.
[33] H. Wang, H. He, Y. Bai, and H. Yue, "Parameterized deep Q-network based energy management with balanced energy economy and battery life for hybrid electric vehicles," *Appl. Energy*, vol. 320, Mar. 2022, Art. no. 119270, doi: 10.1016/j.apenergy.2022.119270.
[34] J. Peng et al., "Multiple electric components health-aware eco-driving strategy for fuel cell hybrid electric vehicle based on soft actor-critic algorithm," *IEEE Trans. Transp. Electrific.*, vol. 10, no. 3, pp. 6242–6257, Sep. 2024, doi: 10.1109/TTE.2023.3339490.
[35] C. Wu, J. Peng, H. He, J. Ruan, J. Chen, and C. Ma, "Health-awareness energy management strategy for battery electric vehicles based on self-attention deep reinforcement learning," *J. Power Sources*, vol. 623, Sep. 2024, Art. no. 235463, doi: 10.1016/j.jpowsour.2024.235463.
[36] S. Hemmati, N. Doshi, D. Hanover, C. Morgan, and M. Shahbakhti, "Integrated cabin heating and powertrain thermal energy management for a connected hybrid electric vehicle," *Appl. Energy*, vol. 283, Nov. 2020, Art. no. 116353, doi: 10.1016/j.apenergy.2020.116353.
[37] H. He, H. Jia, C. Sun, and F. Sun, "Stochastic model predictive control of air conditioning system for electric vehicles: Sensitivity study, comparison, and improvement," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 4179–4189, Sep. 2018, doi: 10.1109/TII.2018.2813315.
[38] Y. C. Park, Y. Kim, and H. Cho, "Thermodynamic analysis on the performance of a variable speed scroll compressor with refrigerant injection," *Int. J. Refrigeration*, vol. 25, pp. 1072–1082, 2002.
[39] Y. Ma, H. Ding, Y. Liu, and J. Gao, "Battery thermal management of intelligent-connected electric vehicles at low temperature based on NMPC," *Energy*, vol. 244, 2022, Art. no. 122571, doi: 10.1016/j.energy.2021.122571.
[40] M. R. Amini, H. Wang, X. Gong, D. Liao-Mcpherson, I. Kolmanovsky, and J. Sun, "Cabin and battery thermal management of connected and automated hevs for improved energy efficiency using hierarchical model predictive control," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 5, pp. 1711–1726, Sep. 2020, doi: 10.1109/TCST.2019.2923792.
[41] C. Wu, J. Ruan, H. Cui, B. Zhang, T. Li, and K. Zhang, "The application of machine learning based energy management strategy in multi-mode plug-in hybrid electric vehicle, part I: Twin delayed deep deterministic policy gradient algorithm design for hybrid mode," *Energy*, vol. 262, 2023, Art. no. 125084, doi: 10.1016/j.energy.2022.125084.
[42] C. Wu et al., "Battery health-considered energy management strategy for a dual-motor two-speed battery electric vehicle based on a hybrid soft actor-critic algorithm with memory function," *Appl. Energy*, vol. 376, 2024, Art. no. 124306, doi: 10.1016/j.apenergy.2024.124306.
[43] H. Xiao, L. Fu, C. Shang, X. Bao, X. Xu, and W. Guo, "Ship energy scheduling with DQN-CE algorithm combining bi-directional LSTM and attention mechanism," *Appl. Energy*, vol. 347, Jun. 2023, Art. no. 121378, doi: 10.1016/j.apenergy.2023.121378.
[44] L. Yu, S. Huo, Z. Wang, and K. Li, "Hybrid attention-oriented experience replay for deep reinforcement learning and its application to a multi-robot cooperative hunting problem," *Neurocomputing*, vol. 523, pp. 44–57, 2023, doi: 10.1016/j.neucom.2022.12.020.
[45] L. Han, K. Yang, T. Ma, N. Yang, H. Liu, and L. Guo, "Battery life constrained real-time energy management strategy for hybrid electric vehicles based on reinforcement learning," *Energy*, vol. 259, Apr. 2022, Art. no. 124986, doi: 10.1016/j.energy.2022.124986.

**Changcheng Wu** received the M.S. degree in mechanical engineering from Beijing University of Technology, Beijing, China, in 2023. He is currently working toward the Ph.D. degree in transportation with the Vehicle-City Integration and Innovation Research Team, Southeast University, Nanjing, China.

His research interests include energy management strategies and thermal management strategies for multipower electric vehicles, intelligent vehicle driving, and deep reinforcement learning for automotive applications.

**Jiankun Peng** received the Ph.D. degree in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2016.

From June 2016 to November 2019, he was a Postdoctoral Researcher with the National Engineering Laboratory of Electric Vehicles, Beijing Institute of Technology, Beijing, China. He is currently an Associate Professor with the School of Transportation, Southeast University, Nanjing, China. He has been named to Stanford's top 2% of the world's top scientists for two years in a row (2022 & 2023 & 2024). He has more than 10 years of research and working experience in modeling and control for new energy vehicle, where he has contributed more than 100 papers.

**Dawei Pi** received the Ph.D. degree in vehicle engineering from Southeast University, Nanjing, China, in 2010.

He is currently a Professor with the Safety and Intelligence of Vehicle Equipment Research Center, Nanjing University of Science and Technology, Nanjing, China. His research interests include vehicle dynamics, intelligent safety control, new energy-intelligent vehicle technology, and so on.

**Xin Guo** received the M.S. degree in mechanical engineering from Shandong University of Technology, Shandong, China, in 2012. He is currently working toward the Ph.D. degree in transportation with Southeast University, Nanjing, China.

His current research interests include thermal management strategies for electric vehicles, and deep reinforcement learning for automotive applications, cooperative systems and control, with applications in connected and autonomous vehicles.

**Hailong Zhang** received the M.S. degree in mechanical engineering from the School of Mechanical and Vehicle Engineering, Beijing Institute of Technology, Beijing, China, in 2018, and the Ph.D. degree in transportation from the School of Transportation, Southeast University, Nanjing, China, in 2022.

He is currently an Associate Professor with the School of Mechanical and Electrical Engineering, North University of China. His research interests include intelligent transportation systems, connected and automated vehicle, and energy-saving strategy.

**Zexing Wang** received the Ph.D. degree in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2023.

He is currently with the National New Energy Vehicle Technology Innovation Center, Beijing, China. His research interests include new energy vehicles, plug-in electric vehicles.

**Yunpeng Li** is currently working toward the M.S. degree in transportation with the Vehicle-City Integration and Innovation Research Team, Southeast University.

His research interests include the health management of electric vehicles and the application of deep reinforcement learning in autonomous vehicles.